Building Momentum: Scaling Whole-Building Energy Programs with Smart Models

Cory Read, EcoMetric, Boise, ID
Michael Honeychuck, EcoMetric, Wilmington, DE
Gavin Zheng, IESO, Toronto, ON

ABSTRACT

As utilities transition from traditional lighting-based efficiency programs to more flexible whole-building approaches, models like Pay-for-Performance (P4P), Strategic Energy Management (SEM), and deep retrofit programs are gaining momentum but still often remain on the sidelines of program portfolios. Whole building program designs emphasize actual, metered energy savings and offer participants greater autonomy in selecting measures, scheduling projects, and earning incentives. They also streamline administrative requirements by reducing the need for intrusive data collection and complex engineering assumptions.

However, scaling whole-building programs introduces significant challenges particularly around data acquisition, baseline modeling, and accommodating diverse building types and operational conditions. Custom modeling for each project, while preferrable, becomes potentially unsustainable at scale.

This paper reviews the whole-building landscape, including market benefits and challenges, and then presents lessons-learned strategies for better designing and implementing whole-building programs. The lessons learned are informed by direct experience developing analysis software and multi-year evaluation work with the Independent Electric System Operator's (IESO) Energy Performance Program (EPP).

The strategies presented are intended for evaluators, utility program managers, and implementers seeking to balance cost-effectiveness with flexibility. As the industry shifts toward whole-building approaches, these insights offer practical guidance to maximize impact and support long-term program success.

Introduction

Whole-building energy efficiency savings programs have long served as a way to expand or enhance existing program portfolios particularly for utilities seeking more flexible strategies to meet regulatory requirements or resource planning targets. These program designs offer both simplifications and new challenges for administrators and participants alike.

For this paper, we define whole-building programs as approaches that capture total site-level savings, in contrast to traditional prescriptive programs that average deemed savings for specific technologies or end uses. Custom and deep prescriptive programs that require whole building modeled savings would fall under the paper's findings. Our focus is on program models that use statistical baseline energy models to estimate savings, rather than physics-based simulation tools such as EnergyPlus[1] (Energy 2025), eQUEST[2] (eQuest, 2025), IES VE[3] (ies 2025), or TRNSYS[4] (TRNSYS 2025), though many of the best practices discussed are broadly applicable across both approaches.

Drawing from evaluation experience, this paper briefly reviews the current program landscape and shares key best practices related to program design, whole-building modeling, customer participation,

[1] https://energyplus.net/
[2] https://www.doe2.com/equest/
[3] https://www.iesve.com/software/virtual-environment
[4] https://www.trnsys.com/

and lessons learned. The primary emphasis is on modeling strategies, with the goal of offering quick-start guidance for utilities exploring the addition of whole-building programs to their efficiency portfolios.

Whole-Building Program Landscape

Whole-building efficiency programs and savings protocols have been in use for many years, but for many utilities, whole-building approaches have often remained peripheral programs, typically limited to pilot initiatives or smaller-scale efforts within broader portfolios. Table 1 lists and describes examples of current whole-building program designs along with their differing characteristics from other programs.

Table 1. Common program designs relying on whole-building savings approaches

Program Type	Key Features	Differentiators
SEM	Training + operational focus	Long-term engagement, behavioral + operational savings
Pay-for-Performance (P4P)	Incentives for actual savings	Measure-agnostic, tied to metered performance
Custom Retrofit	Project-specific engineering	Tailored to complex sites, flexible scope
Deep Energy Retrofit	Large, integrated upgrades	High savings thresholds, often capital-intensive
Prescriptive Bundles	Standard measure packages	Easier path for smaller sites
Retro commissioning (RCx)	Tune-ups + M&V	Low-cost improvements, system optimization

Pay for performance (P4P) is one of the more common program design utilizing whole building savings approaches. P4P programs typically involve a baseline energy usage assessment, followed by ongoing monitoring to track reductions achieved through operational improvements, retrofits, or behavioral changes. Payments are often structured as a fixed rate per unit of energy saved (e.g., per kWh or therm), encouraging deeper and more persistent savings. P4P designs also promote flexibility, allowing customers to choose the most cost-effective strategies for their facilities while aligning utility incentives with actual performance outcomes. Table 2 below summarizes prominent P4P program offerings.

Table 2. List of P4P Programs in North America

Program Administrator	Sector	Building Type	Location
Seattle City Light (SCL)	Large Commercial	Existing	Seattle, WA
New Jersey Clean Energy	Small/Medium Commercial	New Construction	New Jersey
DCSEU	Large Commercial	Existing	Washington DC
Puget Sound Energy (PSE)	Commercial	Existing	Washington
IESO	Commercial	Existing	Ontario (Canada)
PG&E	Residential	Existing	California
Efficiency Vermont	Large Commercial	Existing	Vermont
National Grid	Commercial, Residential	Existing	Massachusetts

Program Administrator	Sector	Building Type	Location
Bay Area Regional Energy Network (BayREN)	Commercial	Existing	San Francisco, CA
National Grid / NYSERDA	Residential	Existing	New York

Program Design Benefits and Barriers

Whole-building program designs for the business sector offer several key advantages over standard prescriptive programs, typically built around predefined measures and deemed savings. The following are some common benefits of whole building program designs:

- Measure flexibility – any mix of strategies
- Performance incentives – only paid for measured savings
- Deeper savings – capture integrated measures
- Participation ease – participants choose project types and depth

Despite their advantages, industry barriers have contributed to the slower adoption of whole-building programs. Common reported barriers include:

- Technical – baseline modeling, non-routine energy events
- Market awareness – unfamiliar relative to traditional programs
- Data access – availability of smart meter data, data privacy

Lessons Learned from IESO's Energy Performance Program (EPP)

The previous discussion provided a brief overview of whole-building program types, key benefits, and common barriers. This paper focuses its analysis on lessons learned from evaluating the Independent Electricity System Operator's (IESO) Energy Performance Program (EPP) over three years of its most recent program funding cycle. EPP is most closely aligned with the pay-for-performance (P4P) whole-building program model. The lessons learned are organized into three key areas:

- Program delivery design
- Participation
- Modeling

Savings and participation have increased over the three years that IESO has administered the program during the most recent budget cycle. Participation in the current budget cycle started out small with only three participants in 2022 growing to 75 facilities in 2024. Figure 1 shows the evaluated and verified savings impacts by building type for closed 2024 projects. The majority of savings (36%) came from grocery and food stores, with 43 participating facilities. Industrial facilities accounted for 31% of total savings, despite representing only 5 sites.

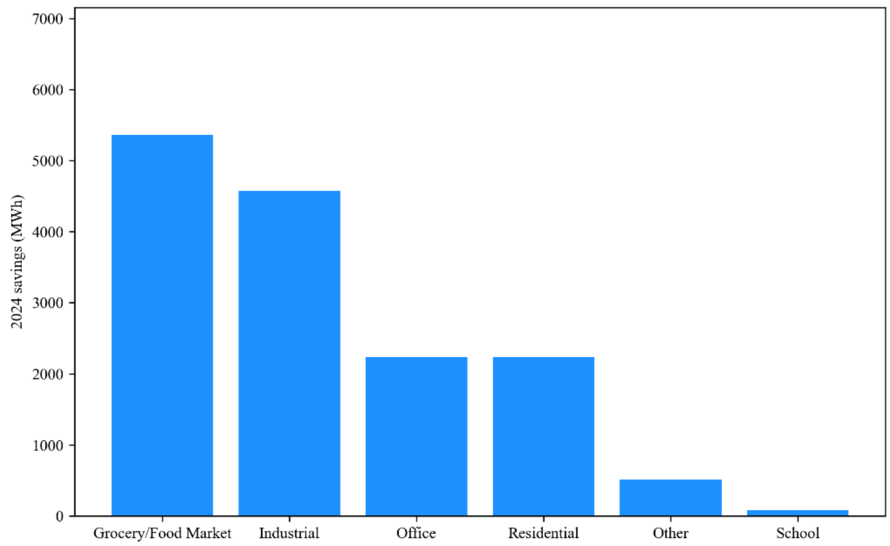

Figure 1. Energy savings by facility types, 2024 program year

Program Delivery Model

IESO transitioned from an Energy Service Provider (ESP) model to a Program Delivery Vendor model to streamline its energy efficiency program. Under the previous model, participants were required to hire ESP consultants to develop baseline energy models and savings reports, which added complexity, costs, and redundant communication cycles. The prior ESP model required the participant, ESPs, third-party technical reviewers (quasi program administrators), and IESO internal program staff to all be involved in final savings and incentive decisions creating potentially unnecessary communication layers.

In contrast, the current Program Delivery Vendor model simplifies participation by centralizing key functions such as data collection, modeling, and savings calculations within a dedicated program delivery vendor. Participants may still choose to engage ESPs, but doing so is now optional, which reduces their overall burden. IESO continues to provide oversight and administrative support under both models, but the new approach enhances efficiency and improves accessibility for participants.

Table 5 presents a side-by-side comparison of the two models, highlighting key differences and similarities.

Table 5. Features of previous and current program delivery models

Feature	IESO/Energy Service Provider Model	IESO/Program Delivery Vendor Model
Program Administration	Fully managed by IESO	Shared between IESO and vendor
Participant Role	May hire an ESP for modeling and/or efficiency project planning or complete with in house facility staff	May optionally hire ESP for project screening/modeling/admin support
Data & Modeling	ESPs or participant develop baseline and savings models	Vendor handles data, modeling, and savings calculations
Technical Review	3rd party reviews ESP/participant models	3rd party creates and produces baseline energy models
Efficiency	Less efficient due to participant burden/communication loop	More streamlined, near turnkey for participant and IESO staff
Final Evaluation	Conducted at program year-end	Conducted at program year-end

Removing the requirement for participants to develop energy models themselves or hire consultants has improved program efficiency. However, the full impact of shifting more responsibility for participant recruitment to the Program Delivery Vendor remains to be seen.

Participation

Successful recruitment for whole-building energy programs often depends on identifying customers who can enroll multiple facilities such as national or regional chains and often act as anchor clients to support early program ramp up. Anchor participants help scale the program quickly and demonstrate early success. IESO has an anchor client enrolling multiple facilities for their EPP program in both the current and previous program budget cycles. However, in jurisdictions where cost-effectiveness is adjusted for free ridership, anchor clients may be flagged as likely free riders due to their existing energy management creating cost-effectiveness issues. Free ridership risks will diminish as the program grows and broadens market engagement, reducing the influence of any single participant.

Another effective strategy targets clusters of customers with similar building types, such as grocery stores, office buildings, or high-rise multifamily properties. These sectors typically share repeatable load profiles, operational patterns, and savings opportunities, making it easier to scale modeling, outreach, and incentive design across multiple sites. Two separate grocery chains have been active in EPP, submitting multiple program facilities.

Figure 2 shows the distribution of building types for projects closed out and paid incentives in 2024. The vast majority of participating facilities were grocery stores, mainly from two separate companies spanning multiple brands enrolling multiple locations.

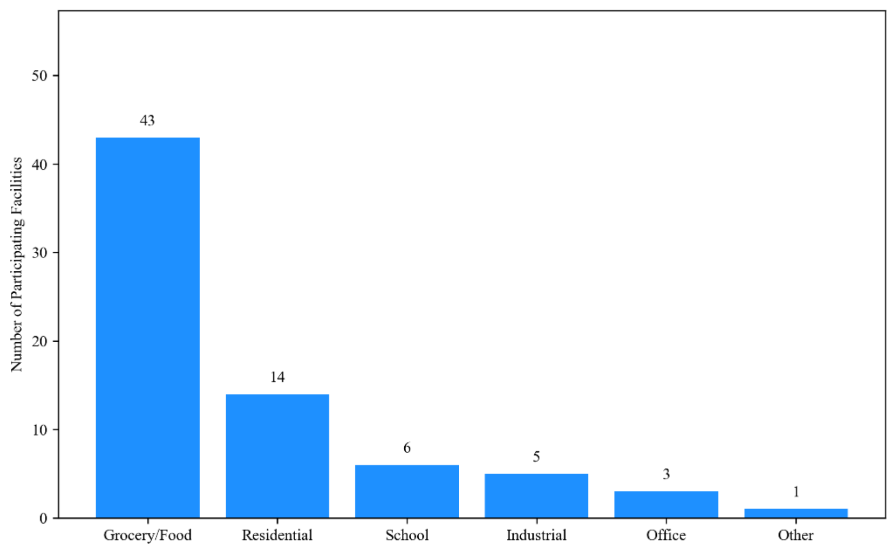

Figure 2. Distribution of EPP participants by building type, 2024 program year

Recruiting industrial customers introduces additional complexity. Facilities with energy use driven by specific production processes rather than weather are harder to model using standard whole-building approaches. Savings often occur in discrete end uses that whole-building meter data may not capture, requiring highly customized models for each site. Consequently, industrial participants may need a separate program track, customized modeling approaches, or process-level submetering to ensure accurate savings attribution.

Developing a Core Model Structure

Weather focused baseline energy models have historically provided a foundation for whole-building energy analysis, but relying solely on weather variables, as in traditional methods like PRISM[5] (PRISM 2025)(optimized degree day models) throughout the 1980's through 2000's or current attempts to standardize modeling such as CalTRACK[6] (CalTRACK 2025) can limit their effectiveness for modern program needs. At the other extreme, fully automated software or deep learning or other statistical learning models often exceed what is necessary for typical 8,760-hour building datasets. The most effective modeling approaches balance complexity and interpretability, capturing the key drivers of energy use without adding unnecessary computational overhead. Ordinary least squares regression (OLS) models remain a reliable choice when using available data. OLS is typically relatively unbiased especially when compared to more complex machine learning or statistical learning algorithms but may have higher variance than more complex methods. ASHRAE (Haberle 2023) whole-building modeling standards have long emphasized the importance of following "best statistical practices" to ensure model quality and validity.

A consistent, well-selected set of input variables can support accurate and scalable modeling across different building types. These variables typically include temperature, temporal (e.g. weekday, hour of the day, month), occupancy schedules, operating hours, and calendar flags such as holidays or weekends. Most buildings can be modeled effectively using a common structure with minimal or no site-specific adjustments. This consistency simplifies implementation and interpretation at scale.

Table 6 presents the core variables used for the majority of savings verifications across two evaluation cycles and nearly 100 fully modeled projects. Only a small subset of industrial projects consistently deviated from the standard model format.

Table 6. Core variables used to baseline model most IESO evaluated facilities

Variable	# Inputs	Description
Degree days/hours	2	Optimized temperature setpoints for a specific site
Holidays	1	Holiday indicator
Temperature splines	6-8	Linear temperature spline buckets. Helps OLS model approximate linear relationship between energy use and temperature
Month	11	Month indicators. Key for minimizing model residuals during peak periods
Peak periods	1	Indicator identifying peak demand time period
Time of week	167	Indicator for hour of the week. Approximates operating schedules. Creates potentially over-parameterized model but works. May not be useful if only partial year data available.
Production/Occupancy	1	Participant provided occupancy schedules and/or any form of production data correlated with baseline energy usage results in a better baseline model
Squared temperature*	1	Occasionally was the best temperature variable
Day of week/weekend*	2-6	Sometimes a day of week or weekend model shift captured dynamics outside of time of week pushes. Could replace time of week approach
Cooling season*	1	AC load bump up indicator occasionally helped. Ex. May thru September.

*Used for some models, but not consistent

[5] https://marean.mycpanel.princeton.edu/images/prism_intro.pdf
[6] https://www.caltrack.org/

The standard model's core variables effectively captured the majority of energy use variance. This approach consistently produced average R-squared values around 90%, indicating strong explanatory power and meeting the program's minimum guideline of 90%. Table 7 presents the average overall R-squared values across different building types.

Key observations related to the recommended variables in Table 6 include that facilities that provide site specific data such as occupancy or production related data will often have better model fits overall. When site specific data is not available variables such as month, time of week, hour of the day, and weekday can approximate building occupancy or operation schedules. Peak summer periods will have increased model accuracy when temporal variables such as month, weekday, or variables identifying utility peak periods are baseline model inputs.

Table 7. Average R-squared values for final evaluated models by building type

Building Type	Average R-squared
Grocery/Food	90%
Residential	90%
Industrial	86%
Other	89%
All Building Types	90%

Figure 3 shows the average evaluated baseline model CV(RMSE) metric by building type. CV(RMSE) is calculated by dividing the standard root mean square error by the average hourly energy usage and is expressed as a percentage. Lower CV(RMSE) values indicate better model accuracy. Among evaluated projects, grocery/food stores achieved the lowest average CV(RMSE) values (4.7%), outperforming both industrial and residential participants. This suggests that, in addition to generating the majority of savings, grocery store savings were generally more reliable.

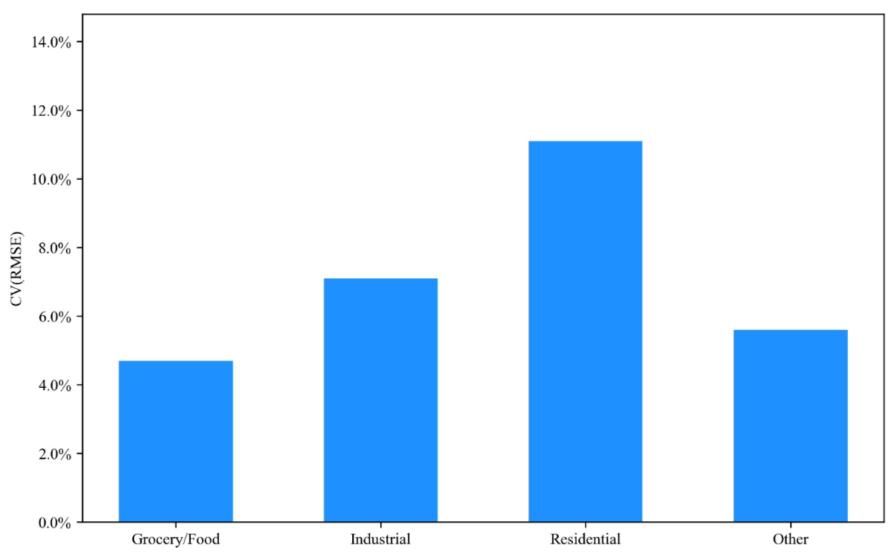

Figure 3. Baseline energy model CV(RMSE) average values by building type

Model Management Challenges

With abundant open-source libraries, data management and model construction is simple, transparent, and repeatable using one or more nearly universally available open-source Python and R libraries.

Handling non-routine adjustments (NRAs) such as major operational changes, equipment failures, or unusual occupancy patterns is a key requirement for program-scale modeling. Emerging software platforms are attempting to offer tools to tag, document, and adjust for these events, making it easier to maintain model accuracy without extensive manual review, but for now program implementers may need to be ready to abandon templates or set software routines to tackle an ad hoc baseline model.

Summary

As the energy efficiency landscape evolves, whole-building programs are emerging as a potential alternative and/or addition to traditional prescriptive approaches and program portfolios. This paper has outlined practical strategies for scaling these programs, drawing on real-world insights from the IESO's Energy Performance Program including the following key recommendations:

- Target an anchor participant to help scale new programs.
- Focus on participants that have multiple sites eligible for enrollment.
- Shift baseline model and meter data acquisition to a third-party firm.
- Standardized model formats and input variables will perform well for most participants and create efficiency whether leveraging software or modeling templates.
    - Baseline models with site specific data such as occupancy/operation data tend to have better model fits.
    - Temporal variables (e.g., month, weekday, hour of day) can help approximate production/occupancy impacts in the absence of site-specific data.
    - Temporal model inputs also help improve model accuracy during peak summer periods.
- Grocery/food stores are good potential participants with low-variance baseline models and consistent savings.
- Industrial participants pose challenges - often requiring custom model building and more time for approving final savings - but they can provide high savings.
- Leveraging scalable tools, including readily available open source analysis libraries, utilities can overcome common technical program implementation barriers.

Special thanks to the IESO program staff for their collaboration and for sharing these key findings with the broader evaluation industry.

Sources

"EnergyPlus." EnergyPlus Energy Simulation Software. U.S. Department of Energy, https://energyplus.net. Accessed 27 June 2025.

"eQUEST – DOE-2 eQUipment Simulation Tool." eQUEST, Architectural Energy Corporation, https://www.doe2.com/equest/. Accessed 27 June 2025.

"IES Virtual Environment | Design Sustainably with Ease." IES Virtual Environment, Integrated Environmental Solutions, https://www.iesve.com/software/virtual-environment. Accessed 27 June 2025.

"TRNSYS – Transient System Simulation Tool." TRNSYS Official Website, University of Wisconsin–Madison, https://www.trnsys.com. Accessed 27 June 2025.

"PRISM - An Introduction." Princeton Scorekeeping Method (PRISM), Princeton University, https://marean.mycpanel.princeton.edu/images/prism_intro.pdf. Accessed 27 June 2025.

"Evaluation, Measurement and Verification." Sector Participants – Energy Efficiency, Independent Electricity System Operator, https://www.ieso.ca/en/Sector-Participants/Energy-Efficiency/Evaluation-Measurement-and-Verification. Accessed 27 June 2025.

"PY2023 Energy Performance Program Evaluation Report." 2021–2024 CDM Framework, Independent Electricity System Operator, 12 July 2024, https://ieso.ca/-/media/Files/IESO/Document-Library/conservation/EMV/2023/PY2023_2021-2024-CDM-Framework_Energy-Performance-Program_Evaluation-Report.pdf. Accessed 27 June 2025.

"Energy Performance Program." Save on Energy: For Business and Industry, Independent Electricity System Operator, n.d., www.saveonenergy.ca/For-Business-and-Industry/Programs-and-incentives/Energy-Performance-Program. Accessed 27 June 2025.

"Pay for Performance Portal." New Jersey's Clean Energy Program, New Jersey Board of Public Utilities Office of Clean Energy, https://www.njcleanenergy.com/residential/prp-portal. Accessed 27 June 2025.

"CalTRACK." CalTRACK Technical Documentation, LF Energy (OpenEEmeter Project), https://www.caltrack.org. Accessed 27 June 2025.

Haberl, J. Dennis, et al. "What Is Contained in Guideline 14-2023." ASHRAE Journal, Feb. 2023, American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), https://www.ashrae.org/file%20library/technical%20resources/ashrae%20journal/haberl--additional-material--feb2023.pdf. Accessed 27 June 2025.

"IPMVP." Protocols, Efficiency Valuation Organization, https://www.evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp. Accessed 27 June 2025.

Agnew, Ken, and Mimi Goldberg. "Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol." The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures, National Renewable Energy Laboratory, Nov. 2017, https://docs.nrel.gov/docs/fy17osti/68564.pdf. Accessed 27 June 2025.